

Sup-norms of Eigenfunctions on Arithmetic Ellipsoids

Valentin Blomer¹ and Philippe Michel²

¹Mathematisches Institut, Universität Göttingen, Bunsenstr. 3-5, 37073 Göttingen, Germany and ²EPFL/SB/IMB/TAN, Station 8, CH-1015 Lausanne, Switzerland

Correspondence to be sent to: vblomer@math.toronto.edu

Let B be a positive quaternion algebra, and let $\mathcal{O} \subseteq B$ be an Eichler order. There is associated, in a natural way, a variety $X = X(\mathcal{O})$ the connected components of which are indexed by the ideal classes of \mathcal{O} and are isomorphic to spheres. This variety is naturally equipped with a Laplace operator and a large family of Hecke operators. For a joint eigenfunction φ of the Hecke algebra and of the Laplace operator with eigenvalue λ , the hybrid sup norm bound $\|\varphi\|_\infty \ll (tV)^{-\delta} t^{1/2} \|\varphi\|_2$ for any $\delta < \frac{1}{60}$ is shown, where $t = (1 + \lambda)^{1/2}$ and $V = \text{vol}(X(\mathcal{O}))$.

1 Introduction

1.1 General setup

Let X be a compact Riemannian manifold and let Δ denote the Laplace operator. A central issue is the study of the behavior of Laplace eigenfunctions

$$\Delta\varphi + \lambda\varphi = 0,$$

as $\lambda \rightarrow +\infty$; in particular, a classical problem is to estimate the L^∞ norm $\|\varphi\|_\infty$. We refer to [18] for a very general and enlightening description of the latter problem along with very precise conjectures about what to expect. For instance, when X is two-dimensional,

Received September 30, 2010; Revised November 18, 2010; Accepted November 30, 2010
 Communicated by Prof. Peter Sarnak

the following standard bound holds:

$$\|\varphi\|_\infty \ll_X (1 + \lambda)^{1/4} \|\varphi\|_2. \quad (1.1)$$

This bound is essentially sharp, as is seen by considering $X = S^2$ and φ a zonal spherical function (the restriction to S^2 of a harmonic homogeneous polynomial in \mathbb{R}^3 which is invariant by the group of rotations fixing a given point on S^2). On the other hand, on hyperbolic surfaces stronger bounds are expected as an effect of negative curvature. These expectations are supported by the groundbreaking work of Iwaniec and Sarnak [11], who proved the following bound for X a Riemann surface coming from an indefinite quaternion algebra over \mathbb{Q} (a Shimura curve) and φ a Hecke–Laplace eigenform:

$$\|\varphi\|_\infty \ll_X (1 + \lambda)^{\frac{1}{4} - \frac{1}{24} + \varepsilon} \|\varphi\|_2. \quad (1.2)$$

When X is compact, the above bound is conjectured to hold with $\frac{1}{4} - \frac{1}{24}$ replaced by 0.

The discrepancy between the positive curvature case (say the sphere) and the hyperbolic case is closely related to the fact that in the former, Laplace eigenvalues have high multiplicities while in the latter, the multiplicities are expected to be essentially bounded. Very little is known about these multiplicities in the hyperbolic case (even for arithmetic surfaces), and this explains why Iwaniec and Sarnak considered Hecke eigenforms for which multiplicity one theorems hold (of course if multiplicities are very small, as expected, restricting to Hecke eigenforms does not reduce the generality).

In this paper, we reconsider and extend the previous reasoning to the 2-sphere which may be realized as a connected component of a locally homogeneous space of arithmetical type associated with a definite quaternion algebra defined over \mathbb{Q} . It is therefore equipped with a large commutative ring of “Hecke operators” commuting with Δ . Considering Laplace eigenfunctions which are also eigenfunctions of the Hecke operators eliminates the issue of high multiplicities, and we can obtain a bound analogous to (1.2) (with $\frac{1}{24}$ replaced by $\frac{1}{48}$, cf. (5.13)).

1.2 Hybrid bounds

We should point out, however, that our main interest here is somewhat different from the setup in [11] as we are interested in how (1.1) depends on X . This question was investigated in great generality in [4] where it is shown that the dependency involves

the sectional curvature and the injectivity radius. More precise results are available for locally symmetric spaces $\Gamma \backslash G/K$. For suitable families of hyperbolic surfaces $X_\Gamma = \Gamma \backslash \mathbb{H}$ (including suitable coverings of a fixed surface [14, 17] or arithmetic surfaces), the bound (1.1) takes the more precise form

$$\|\varphi\|_\infty \ll \text{vol}(X)^{o(1)}(1+\lambda)^{1/4}\|\varphi\|_2, \quad (1.3)$$

where the metric on X_Γ is the one descending from a fixed hyperbolic metric on \mathbb{H} . It is natural to try to improve on (1.3) simultaneously in the eigenvalue and the “volume” aspect. Such bounds are called “hybrid”. The first example of such a hybrid bound was obtained in the work of Holowinsky and the Blomer [2]: as X_Γ varies over the sequence of (noncompact but finite volume) modular curves of level square-free level N ,

$$X_0(N) = \Gamma_0(N) \backslash \mathbb{H}^* \rightarrow X_0(1) = \text{SL}_2(\mathbb{Z}) \backslash \mathbb{H}^*, \quad N \text{ square-free},$$

one has, for any Hecke–Maass cuspform φ on $X_0(N)$,

$$\|\varphi\|_\infty \ll ((1+\lambda)^{1/2} \text{vol}(X_0(N)))^{-\delta} (1+\lambda)^{1/4} \|\varphi\|_2, \quad (1.4)$$

where $\delta > 0$ is an absolute constant (let us recall that $\text{vol}(X_0(N)) = N^{1+o(1)}$). The best possible bound in this situation would be $\delta = \frac{1}{2} + o(1)$ which among other things would imply the Lindelöf hypothesis for the Hecke L -function $L(\varphi, s)$ (see Section 6 for a related discussion).

In this paper, we obtain a hybrid analog of (1.4) for 2-spheres. As we explain in Section 4 (see, in particular (4.1)), to a definite quaternion algebra B defined over \mathbb{Q} and an order $\mathcal{O} \subset B(\mathbb{Q})$, there is associated a finite disjoint union of (finite quotients of) spheres S^2 :

$$X := X(\mathcal{O}) = \bigsqcup_{[I] \in \text{Cl}(\mathcal{O})} X_{[I]}, \quad X_{[I]} := \Gamma_{[I]} \backslash S^2,$$

$$\Gamma_{[I]} < \text{SO}_3(\mathbb{R}) \text{ some finite subgroup of order } \leq 12.$$

The quotients $X_{[I]}$ are the *components* of X and the (finite, but possibly large) set indexing the components is the set of *classes* of \mathcal{O} -ideals.

The variety X is equipped with a “natural” Riemannian metric and a corresponding volume form: the metric is obtained by choosing once and for all a left $\mathrm{SO}_3(\mathbb{R})$ -invariant Riemannian metric on S^2 and by putting on each component $X_{[I]}$ the induced metric multiplied by $1/|\Gamma_{[I]}|$; in this way, $\mathrm{vol}(X_{[I]}) \asymp 1$, and the volume of X has roughly the size of the class group:

$$\mathrm{vol}(X(\mathcal{O})) \asymp |\mathrm{Cl}(\mathcal{O})| = \mathrm{disc}(\mathcal{O})^{1/2+o(1)},$$

cf. (2.4). As a Riemannian manifold, X is equipped with a Laplace operator Δ . Moreover, being arithmetically defined, X is also endowed with a commutative algebra of Hecke correspondences \mathbf{T} which is generated by Hecke correspondences $(T_p)_p$, indexed by the primes p coprime with $\mathrm{disc}(\mathcal{O})$. Each correspondence T_p is of degree $p+1$ and yields a self-adjoint Hecke operator (w.r.t. the measure derived from the Riemannian metric) commuting with Δ . Our main result is the following:

Theorem 1. Let $X(\mathcal{O})$ be as above and suppose that \mathcal{O} is an Eichler order. Let φ be a Laplace eigenform φ with eigenvalue λ which is also an eigenform of the Hecke algebra \mathbf{T} . One has

$$\|\varphi\|_\infty \ll ((1+\lambda)^{1/2} \mathrm{vol}(X(\mathcal{O})))^{-\delta} (1+\lambda)^{\frac{1}{4}} \|\varphi\|_2$$

for some absolute constant $\delta > 0$. Any value $\delta < \frac{1}{60}$ is admissible. \square

Remark 1. The result is nontrivial already for the trivial eigenvalue $\lambda = 0$. In this case, φ is constant on each of the components $X_{[I]}$. The space of locally constant functions has a very deep number theoretic meaning. For instance, if B is the definite quaternion algebra ramified at a single prime q and \mathcal{O} is a maximal order, this space is identified with the set of function on the set of isomorphy classes of supersingular elliptic curves in characteristic q (see [8, Section 2]). The space of such functions has large dimension, equal to $|\mathrm{Cl}(\mathcal{O})| \approx \mathrm{vol}(X(\mathcal{O}))$. Again we see that (1.3) is in general sharp in the volume aspect: pick the function which is 1 on a given component and 0 on the others (a similar reasoning involving harmonic polynomials instead of constant functions shows that (1.3) is sharp in both aspects). However, the restriction to Hecke eigenfunction resolves this multiplicity issue, and the meaning of

$$\|\varphi\|_\infty \ll \mathrm{vol}(X(\mathcal{O}))^{-\delta} \|\varphi\|_2$$

is an equidistribution statement for Hecke correspondences: a locally constant function which is moreover a Hecke-eigenfunction cannot accumulate too much of its mass on a single component. \square

Remark 2. There are a priori two main “directions” in which $X = X(\mathcal{O})$ could vary: either by varying the order \mathcal{O} inside B (the main numerical parameter measuring the variation of the Eichler order \mathcal{O} is its level N which is a positive integer coprime with the discriminant of B) or by varying B (which is equivalent to letting $\text{disc}(B) \rightarrow \infty$). Our result is uniform in *both* of these directions of variation. In [2], only the \mathcal{O} -direction was considered (the corresponding quaternion algebra in that case is the algebra $B = M_2(\mathbb{Q})$ of 2×2 matrices), although this has now been substantially extended and improved by Templier [21] to cover also the case when B varies amongst the *indefinite* quaternion algebras. \square

Remark 3. The numerical value of δ follows from an inspection of (5.12) and (5.13). Our aim was to show the existence of some $\delta > 0$ with relatively little technology; with more involved estimations, the exponent $\frac{1}{60}$ could be improved. \square

Remark 4. Another motivation to study sup-norms of eigenfunctions comes from a Waldspurger-type formula due to Gross (see [8, 10]) that relates central values of certain Rankin–Selberg L -functions to averages of automorphic forms over CM-points. In this way, the bound of Theorem 1 in the λ -aspect translates into a subconvex bound for certain Rankin–Selberg L -functions $L(\pi_1 \otimes \pi_2, \frac{1}{2})$ in terms of the Archimedean parameter of π_1 . Although there are by now numerically stronger and more general subconvexity results, the method is new and very different from the usual approaches and based only on the arithmetic of quadratic forms. See Section 6 for a more detailed discussion. \square

1.3 Principle of proof

Our proof proceeds roughly as follows. We consider an amplified second moment which we transform by a pre-trace formula into a sum over a sort of automorphic kernel, see (5.3). This starting point is similar in most investigations of sup-norms of eigenfunction on arithmetically defined manifolds, see, for example, [2, 9, 11, 16, 21]. In all cases one encounters eventually an interesting diophantine problem the solution of which is at the heart of the problem. Here, the analysis diverges in all known cases, and depends

in a nontrivial way on the underlying manifold. In our situation, the sum over the automorphic kernel can be expressed in terms of weighted sums of representation numbers of integers by definite quaternionic norm forms, cf. for instance (5.5) and the subsequent remark which interprets this expression as an average over Fourier coefficients of certain theta-series. These sums can be bounded by the classical reduction theory of quadratic forms and methods from diophantine approximation, but the analysis is quite subtle. To get a nontrivial sup-norm bound in the λ -aspect, one has to estimate the number of integral solutions to $Q(\mathbf{x}) = \ell$ close to two given orthogonal hyperplanes for certain quaternary quadratic forms Q . To get a nontrivial sup-norm bound in the volume aspect, one has to estimate the number of solutions to $Q(\mathbf{x}) = \ell$ for certain quaternary forms Q of very large discriminant (compared with ℓ). These two bounds are given in Lemmas 1 and 2 that appear to be new and may also be useful in other contexts. These two results can be combined to yield a simultaneous bound in both aspects.

The article is organized as follows: Section 2 compiles and recalls basic facts on quaternion algebras. Section 3 is the diophantine heart of the paper and contains bounds for representation numbers of quadratic forms that are needed later. Section 4 realizes the manifold X as an adelic quotient of a quaternion algebra and introduces the relevant operators on this space. In Section 5, we construct a suitable test function and a suitable amplifier for the pre-trace formula and estimate the resulting expression. Section 6 provides a link to the subconvexity problem.

1.4 Possible extensions

The results of this paper can be generalized in various directions. Firstly, a Laplace eigenfunction on S^2 may be viewed as a function on the group $\mathrm{SO}_3(\mathbb{R})$, left-invariant under some suitably embedded $\mathrm{SO}_2(\mathbb{R})$ which is an eigenform of the *Casimir* operator; the method presented here extends readily and provides nontrivial bounds for the more general Hecke–Casimir eigenfunctions. Less evidently perhaps, these methods extend to Hecke–Laplace (and Hecke–Casimir) eigenfunctions associated with totally definite quaternion algebras defined over a fixed *totally real number field* (the associated symmetric space is then a union of quotients of products of spheres.) A nice consequence of this extension to more general number fields is that such bounds provide nontrivial bounds for Hecke–Laplace (more generally Hecke–Casimir) eigenfunctions on certain unions of three-dimensional ellipsoids (more generally orthogonal groups) associated with definite quaternary quadratic forms over \mathbb{Q} . For instance, the result of the present paper essentially provides such bounds when the discriminant of the quaternary form is

a square, while the case of a nonsquare discriminants follows from the above-mentioned generalization to the quadratic extension generated by the square root of the discriminant. Such extensions will be discussed in a subsequent paper.

2 Arithmetic in Quaternion Orders

2.1 Quaternion algebras

We recall some basic terminology and facts of quaternion algebras, see, for example, [23]. For $a, b \in \mathbb{Q}^\times$, let $B = \left(\frac{a,b}{\mathbb{Q}}\right)$ be the corresponding quaternion algebra over \mathbb{Q} , that is,

$$B(\mathbb{Q}) = \mathbb{Q} + \mathbb{Q}i + \mathbb{Q}j + \mathbb{Q}ij$$

with $i^2 = a$, $j^2 = b$, $ij + ji = 0$. Its center $Z(B)$ is the algebra of scalars \mathbb{Q} . If $\gamma = x_1 + x_2i + x_3j + x_4ij \in B$, let $\bar{\gamma} := x_1 - x_2i - x_3j - x_4ij$ be the canonical involution and let tr and nr be the reduced norm and trace

$$\text{tr}(\gamma) = \gamma + \bar{\gamma} = 2x_1, \quad \text{nr}(\gamma) = \gamma\bar{\gamma} = \frac{1}{2}\text{tr}(\gamma\bar{\gamma}) = x_1^2 - ax_2^2 - bx_3^2 + abx_4^2.$$

A place v is called ramified if $B_v := B \otimes_{\mathbb{Q}} \mathbb{Q}_v$ is a division algebra, and nonramified otherwise; in the former case, B_v is the unique (up to isomorphism) quaternion division algebra over \mathbb{Q}_v ; in the latter $B_v \cong \text{Mat}(2, \mathbb{Q}_v)$ in which case the reduced norm and reduced trace are given by the usual determinant and trace and for matrices. We recall that a quaternion algebra is ramified at an even finite number of places and that the finite ramified places divide $2ab$. For the rest of this paper, we assume that $a, b < 0$ so that B is ramified at ∞ (i.e., $B(\mathbb{R})$ is the algebra of real Hamilton quaternions). The *reduced discriminant* D_B of B is the product of the finite ramified primes.

2.2 Lattices, orders

A lattice or ideal $\mathfrak{I} \subset B(\mathbb{Q})$ is a \mathbb{Z} -module of maximal rank 4. The product of two lattices \mathfrak{I}_1 and \mathfrak{I}_2 is given by $\mathfrak{I}_1\mathfrak{I}_2 := \{\gamma_1\gamma_2 \mid \gamma_1 \in \mathfrak{I}_1, \gamma_2 \in \mathfrak{I}_2\}$, and the inverse is given by $\mathfrak{I}^{-1} := \{\gamma \in B \mid \gamma\mathfrak{I} \subseteq \mathfrak{I}\}$. An order \mathcal{O} is a subring of $B(\mathbb{Q})$ which is also lattice, in particular its elements are integral (i.e., $\text{tr}(\gamma), \text{nr}(\gamma) \in \mathbb{Z}$ for $\gamma \in \mathcal{O}$). It follows that an element $\gamma \in \mathcal{O}$ is invertible if and only if $\text{nr}(\gamma) = 1$, and one can show $|\mathcal{O}^\times| \leq 24$ (see [23, p. 145]).

The left (resp. right) order $\mathcal{O}_l(\mathfrak{I})$ (resp. $\mathcal{O}_r(\mathfrak{I})$) of a lattice \mathfrak{I} is the set

$$\mathcal{O}_l(\mathfrak{I}) = \{\gamma \in B(\mathbb{Q}), \gamma\mathfrak{I} \subset \mathfrak{I}\} \quad (\text{resp. } \mathcal{O}_r(\mathfrak{I}) = \{\gamma \in B(\mathbb{Q}), \mathfrak{I}\gamma \subset \mathfrak{I}\}).$$

The left (resp. right) order of a lattice is an order. Given an order \mathcal{O} , a left \mathcal{O} -ideal is defined as a lattice \mathfrak{I} such that $\mathcal{O}_l(\mathfrak{I}) = \mathcal{O}$ (right \mathcal{O} -ideals are defined in the obvious way). Two right \mathcal{O} -ideals \mathfrak{I} and \mathfrak{I}' are called equivalent if there exists $\gamma \in B^\times(\mathbb{Q})$ such that $\mathfrak{I}' = \gamma\mathfrak{I}$. The set of such equivalence classes is denoted $\text{Cl}(\mathcal{O})$; this set is finite and its cardinality $|\text{Cl}(\mathcal{O})|$ is the *class number* of \mathcal{O} .

2.3 Localization

Given a lattice \mathfrak{I} , we denote by $\mathfrak{I}_p = \mathfrak{I} \otimes_{\mathbb{Z}} \mathbb{Z}_p$ the closure of \mathfrak{I} inside B_p .

The choice of a maximal order $\mathcal{O}_{\max} \subset B(\mathbb{Q})$ and of a \mathbb{Z} -basis of \mathcal{O}_{\max} determines, for every unramified p (i.e., not dividing D_B), isomorphisms

$$\mathcal{O}_{\max,p} \simeq M_2(\mathbb{Z}_p) \quad \text{hence of } \mathbb{Q}_p\text{-algebras } B_p \simeq M_2(\mathbb{Q}_p). \quad (2.1)$$

We fix such a choice and in the sequel identify freely elements of B_p (for p unramified) with matrix elements. On the other hand, if p is ramified, then

$$\mathcal{O}_{\max,p} = \{\gamma_p \in B_p, \text{nr}(\gamma_p) \in \mathbb{Z}_p\}$$

is the unique maximal order of B_p .

Two right \mathcal{O} -ideals $\mathfrak{I}, \mathfrak{I}'$ are called *everywhere locally equivalent* if for every prime p there is $\gamma_p \in B_p^\times$ such that $\mathfrak{I}'_p = \gamma_p \mathfrak{I}_p$. This equivalence relation is coarser than the previous one, and having everywhere only one local equivalence class for a given order \mathcal{O} is equivalent to saying that every right \mathcal{O} -ideal is everywhere locally principal.

2.4 Discriminant and reduced discriminant

One has $2\text{nr}(\gamma) = \text{tr}(\gamma\bar{\gamma})$, so the trace defines a nondegenerate bilinear form

$$(\gamma, \gamma') = \frac{1}{2}\text{tr}(\gamma\bar{\gamma}').$$

The discriminant of an order \mathcal{O} is by definition.

$$\mathrm{disc}(\mathcal{O}) = \det(\mathrm{tr}(\gamma_i \bar{\gamma}_j)_{i,j \leq 4})$$

for $\{\gamma_1, \dots, \gamma_4\}$ a \mathbb{Z} -basis of \mathcal{O} . This does not depend on the choice of the basis. (Often the discriminant of \mathcal{O} is defined as $\det(\mathrm{tr}(\gamma_i \gamma_j)_{i,j \leq 4})$ which differs from our definition by a sign.)

The reduced norm $\mathrm{nr}(\mathfrak{I}) \in \mathbb{Q}$ of a lattice \mathfrak{I} is the positive generator of the fractional \mathbb{Z} -ideal generated by all elements $\mathrm{nr}(\gamma)$ with $\gamma \in \mathfrak{I}$. The reduced norm is multiplicative on ideals [23, p. 24].

The dual of a lattice \mathfrak{I} is the lattice

$$\mathfrak{I}^* = \{\gamma \in B(\mathbb{Q}), \mathrm{tr}(\gamma \mathfrak{I}) \subset \mathbb{Z}\}.$$

The *different* (or complement in Eichler's terminology) of an order \mathcal{O} is the dual \mathcal{O}^* of \mathcal{O} . This is an \mathcal{O} -module and one defines the *reduced discriminant* $\mathrm{disc}^*(\mathcal{O})$ of \mathcal{O} to be

$$\mathrm{disc}^*(\mathcal{O}) := \mathrm{nr}((\mathcal{O}^*)^{-1}) = \mathrm{nr}(\mathcal{O}^*)^{-1} \quad (2.2)$$

(by multiplicativity of the norm). If \mathcal{O} is a maximal order, its reduced discriminant equals the reduced discriminant D_B of B defined at the beginning of this section [23, II.4.7].

One has the following relation between discriminant and reduced discriminant [23, I.4.7].

$$\mathrm{disc}(\mathcal{O}) = \mathrm{disc}^*(\mathcal{O})^2. \quad (2.3)$$

(Strictly speaking, the proof in [23] is only carried out in the class number one case and left as an exercise in the general case. In this paper, we consider Eichler orders (see below) and this special case is already enough, since Eichler orders have locally class number one, and both sides of (2.3) are the products of their local components. If the reduced discriminant is square-free, (2.3) is easy to see anyway, for example, [5, p. 131, first paragraph].) All these invariants admit of course local counterparts and the global ones are obtained as products of the local ones.

2.5 Eichler orders

An Eichler order is by definition the intersection of two maximal orders. To an Eichler order there is associated an integer N coprime to D_B such that for every p coprime with D_B , \mathcal{O}_p is conjugate to the order $\left\{ \begin{pmatrix} \mathbb{Z}_p & \mathbb{Z}_p \\ N\mathbb{Z}_p & \mathbb{Z}_p \end{pmatrix} \right\}$ of $M_2(\mathbb{Q}_p)$. (In [23] and many other places this integer is called the “level” of the Eichler order. Eichler [5, Section 2] uses the word *level* for what we called the reduced discriminant. To avoid confusion, we will not use the word *level* for an order, but only for the associated norm form, and we will use instead the terminology “an Eichler order associated with the integer N ”.)

We may and will assume that the choice of \mathcal{O}_{\max} is such that \mathcal{O}_p corresponds precisely to $\left\{ \begin{pmatrix} \mathbb{Z}_p & \mathbb{Z}_p \\ N\mathbb{Z}_p & \mathbb{Z}_p \end{pmatrix} \right\}$ under the identification (2.1). Note also that since \mathcal{O}_p is the unique maximal order of B_p at a ramified prime p , Eichler orders associated with the same N are locally conjugate (and conversely). Finally, since right (resp. left) ideals of an Eichler order \mathcal{O} are locally principal (that is, for every p there is $\gamma_p \in B^\times(\mathbb{Q}_p)$ such that $\mathfrak{I}_p = \gamma_p \mathcal{O}_p$, cf. [23, Section 2]), the left order \mathcal{O}_l of a right \mathcal{O} -ideal is everywhere locally conjugate to \mathcal{O} (that is, $(\mathcal{O}_l)_p = \gamma_p \mathcal{O}_p \gamma_p^{-1}$). In particular, \mathcal{O}_l is conjugate to \mathcal{O} by an element of $B^\times(\mathbb{Q})$.

For an Eichler order \mathcal{O} , the discriminant and the reduced discriminant have the following explicit expressions [23, p. 85]:

$$\text{disc}^*(\mathcal{O}) = D_B N, \quad \text{disc}(\mathcal{O}) = (D_B N)^2,$$

and the class number equals [Vi, p. 143]

$$|\text{Cl}(\mathcal{O})| \asymp D_B N \prod_{p|D_B} \left(1 - \frac{1}{p}\right) \prod_{p|N} \left(1 - \frac{1}{p}\right)^{-1} = \text{disc}(\mathcal{O})^{1/2+o(1)}. \quad (2.4)$$

3 Representation Numbers of Quadratic Forms

In this section, we provide the necessary tools to treat the diophantine problems mentioned at the end of the introduction. We recall some facts about positive definite quadratic forms. Let

$$Q(\mathbf{x}) = \frac{1}{2} \sum_{1 \leq i, j, \leq n} a_{ij} x_i x_j, \quad a_{ij} = a_{ji} \in \mathbb{Z}, \quad a_{jj} \in 2\mathbb{Z},$$

be a positive definite integral quadratic form in n variables. Let $A = (a_{ij})_{1 \leq i, j \leq n}$ be the integral, even (i.e., the diagonal elements are even), symmetric $n \times n$ -matrix associated

with Q . The *determinant* of Q is just

$$\Delta = \det A,$$

and the *level* of Q is the smallest integer N such that NA^{-1} is an integral even matrix. Both the determinant and the level are the products of their local components (i.e., the determinant and the level of the quadratic lattices (\mathbb{Z}_p^n, Q) for p varying over the primes), and hence are same for two quadratic forms that are everywhere locally equivalent.

3.1 Reduction theory

A form Q is called Minkowski-reduced if for all $j = 1, \dots, n$ and for all $\mathbf{x} \in \mathbb{Z}^n$ such that $(\mathbf{e}_1, \dots, \mathbf{e}_j, \mathbf{x})$ can be extended to an integral basis of \mathbb{Z}^n , one has

$$Q(\mathbf{x}) \geq Q(\mathbf{e}_j),$$

cf. [3, chapter 12]. Every Q is \mathbb{Z} -equivalent to a Minkowski-reduced form [3, Theorem 12.1.1]. It is not hard to see [3, Lemma 12.1.1] that a Minkowski-reduced form satisfies

$$0 < a_{11} \leq a_{22} \leq \dots \leq a_{nn}, \quad |2a_{ji}| \leq a_{jj} \quad \text{for } 1 \leq j < i \leq n. \quad (3.1)$$

We can write a Minkowski-reduced form Q as

$$\begin{aligned} Q(\mathbf{x}) &= \sum_{1 \leq i, j \leq n} a_{ij} x_i x_j \\ &= h_1(x_1 + c_{12}x_2 + \dots + c_{1n}x_n)^2 + h_2(x_2 + c_{23}x_3 + \dots + c_{2n}x_n)^2 + \dots + h_n x_n^2. \end{aligned}$$

Then it is known [3, Theorem 12.3.1] that $a_{jj} \asymp h_j$ where the implied constants in the upper and lower bounds depend only on n . Let us define

$$H_j := \min_{i \geq j} h_i \asymp a_{jj}. \quad (3.2)$$

We see inductively that

$$Q(\mathbf{x}) \geq H_j \text{ unless } x_j = x_{j+1} = \dots = x_n = 0. \quad (3.3)$$

We denote generally by

$$0 < \lambda_1 \leq \dots \leq \lambda_n \quad (3.4)$$

the eigenvalues of A . By (3.1) and (3.2) we have

$$\lambda_n \ll a_m \ll H_n \ll \det A. \quad (3.5)$$

By Cramer's rule, the entries of A^{-1} are given by $\det \tilde{A} / \det A \ll 1$ where \tilde{A} is an $(n-1) \times (n-1)$ minor of A . Hence

$$\lambda_1 \gg 1. \quad (3.6)$$

Let μ_n denote the n -dimensional Lebesgue measure. Then

$$\mu_n(\{\mathbf{x} \in \mathbb{R}^n \mid Q(\mathbf{x}) \leq y\}) \asymp_n \frac{y^{n/2}}{\Delta^{1/2}}. \quad (3.7)$$

Given $y > 0$, we denote by $\mu_{n-1}\{\mathbf{x} \in \mathbb{R}^n \mid Q(\mathbf{x}) = y\}$ the measure on the ellipsoid deduced from the restriction of the Euclidean metric on \mathbb{R}^n ; one has

$$\mu_{n-1}(\{\mathbf{x} \in \mathbb{R}^n \mid Q(\mathbf{x}) = y\}) \asymp_n \frac{y^{(n-1)/2}}{(\lambda_1 \cdots \lambda_{n-1})^{1/2}} = \frac{y^{(n-1)/2}}{(\Delta/\lambda_n)^{1/2}} \ll_n \frac{y^{(n-1)/2} H_n^{1/2}}{\Delta^{1/2}}, \quad (3.8)$$

since the axes of the ellipsoid $Q(\mathbf{x}) = y$ have lengths $(y/\lambda_j)^{1/2}$.

We use the notation $r_Q(\ell)$ to denote the number of integral representations of ℓ by Q .

3.2 Diophantine lemmas

This section contains the two diophantine results that will eventually yield nontrivial bounds in the volume aspect and the λ -aspect in Theorem 1.

Lemma 1. Let Q be a positive-definite integral quaternary quadratic form of determinant Δ and level N . For $y \geq 1$, let $\mathcal{A} \subseteq \mathbb{N} \cap [1, y]$ be a subset of integers bounded by y . Then

$$\sum_{\ell \in \mathcal{A}} r_Q(\ell) \ll_\varepsilon \frac{y^2}{\Delta^{1/2}} + \frac{y^{3/2}}{(\Delta/N)^{1/2}} + y^\varepsilon |\mathcal{A}|$$

for any $\varepsilon > 0$, the implied constant depending on ε alone. \square

Proof. We can assume that Q is Minkowski-reduced. Let $\tilde{A} = (a_{ij})_{1 \leq i, j \leq 3}$ be the upper left 3×3 -submatrix of A , and denote by $\tilde{\Delta}$ its determinant. Clearly \tilde{A} is again a

Minkowski-reduced matrix. The entry in the lower right corner of A^{-1} is by Cramer's rule $\tilde{\Delta}/\Delta$; by definition of the level, $N\tilde{\Delta}/\Delta$ must be integral, hence

$$\tilde{\Delta} \geq \frac{\Delta}{N}. \quad (3.9)$$

We distinguish three cases. If $y \geq H_4$, then by (3.7) and (3.8) we have by the Lipschitz principle

$$\sum_{\ell \in \mathcal{A}} r_Q(\ell) \leq \sum_{\ell \leq y} r_Q(\ell) \ll \frac{y^2}{\Delta^{1/2}} + \frac{y^{3/2} H_4^{1/2}}{\Delta^{1/2}} \ll \frac{y^2}{\Delta^{1/2}}.$$

If $H_3 \leq y < H_4$, then by (3.3), all ℓ are represented by the ternary form corresponding to the matrix \tilde{A} , and hence by (3.9) and the same reasoning we find

$$\sum_{\ell \in \mathcal{A}} r_Q(\ell) \ll \frac{y^{3/2}}{\tilde{\Delta}^{1/2}} + \frac{y H_3^{1/2}}{\tilde{\Delta}^{1/2}} \ll \frac{y^{3/2}}{\tilde{\Delta}^{1/2}} \ll \frac{y^{3/2}}{(\Delta/N)^{1/2}}.$$

Finally if $y < H_3$, then again by (3.3) all ℓ are represented by a binary form, so that $r_Q(\ell) \ll y^\varepsilon$. This completes the proof of the lemma. \blacksquare

For the next lemma we introduce the following notation: For a positive-definite integral quaternary quadratic form Q given by a matrix A and some parameter $\xi \geq 0$, we write

$$r_Q(\ell, \xi) := \sup |\{\mathbf{x} \in \mathbb{Z}^4 \mid Q(\mathbf{x}) = \ell, \quad |\mathbf{v}_1^T A \mathbf{x}|^2 + |\mathbf{v}_2^T A \mathbf{x}|^2 \leq \xi\}|$$

where the supremum is taken over all pairs \mathbf{v}_1 and \mathbf{v}_2 of orthonormal vectors in the quadratic space (\mathbb{R}^4, Q) , that is

$$Q(\mathbf{v}_1) = Q(\mathbf{v}_2) = 1, \quad \mathbf{v}_1^T A \mathbf{v}_2 = 0. \quad (3.10)$$

Lemma 2. Let Q be a positive-definite integral quaternary quadratic form of determinant Δ . Let $\ell \in \mathbb{N}$ and $\xi > 0$. Then

$$r_Q(\ell, \xi) \ll_\varepsilon (1 + \xi \Delta)(1 + \Delta^2 \xi + (\Delta^2 \ell^7 \xi)^{1/8})(\ell \Delta (\xi + 1/\xi))^\varepsilon,$$

with an implied constant depending only on ε . Moreover,

$$r_Q(\ell, \xi) \leq r_Q(\ell) \ll_\varepsilon \ell^{1+\varepsilon}. \quad \square$$

Remark. This lemma bounds the number of representations of ℓ by Q the projection to a given 2-dimensional plane of which is small as measured by the parameter ξ . Qualitatively, that lemma says that if ξ is a large enough negative power of $\Delta\ell$, then the number of such representation is essentially bounded. Such qualitative statement is already sufficient to deduce a nontrivial bound for $\|\varphi\|_\infty$ in the λ -aspect. \square

Proof. For small ξ , we need to count lattice points in some slightly thickened S^1 inside S^3 . In other words, we have a slightly perturbed binary problem which explains why the number of representations should be almost bounded. To make this idea precise, we denote the eigenvalues of A as in (3.4) and write $A = B^T B$ for some (unique) positive symmetric matrix B . Let $0 < \mu_1 \leq \dots \leq \mu_4$ be the eigenvalues of B , and write $\|\cdot\|$ for the usual Euclidean 2-norm. If $Q(\mathbf{x}) = \ell$, then $\|A\mathbf{x}\| = \|B^T B\mathbf{x}\| \ll \mu_4 \ell^{1/2} = (\lambda_4 \ell)^{1/2} \ll (\Delta\ell)^{1/2}$ and $\|\mathbf{x}\| \ll (\ell/\lambda_1)^{1/2} \ll \ell^{1/2}$ by (3.5) and (3.6).

Let $V = (v_{ij}) \in \mathbb{R}^{2 \times 4}$, say, be the matrix the two rows of which are given by $\mathbf{v}_1^T A$ and $\mathbf{v}_2^T A$. It is not hard to see that (3.10) implies

$$\sum_{1 \leq j_1 < j_2 \leq 4} \left| \det \begin{pmatrix} v_{1j_1} & v_{1j_2} \\ v_{2j_1} & v_{2j_2} \end{pmatrix} \right|^2 \geq \mu_1 \mu_2 \gg 1.$$

Hence we can pick two columns $j_1, j_2 \in \{1, 2, 3, 4\}$ of V such that the corresponding 2-by-2 matrix V' has a determinant $\gg 1$. Since $v_{ij} \ll \lambda_4^{1/2} \ll \Delta^{1/2}$, we see that the entries of $(V')^{-1}$ are bounded by $O(\Delta^{1/2})$.

Let us denote the other two columns of V by j_3 and j_4 and the corresponding 2-by-2 matrix by V'' . Then $|\mathbf{v}_1^T A\mathbf{x}|^2 + |\mathbf{v}_2^T A\mathbf{x}|^2 \leq \xi$ implies

$$\left\| V' \begin{pmatrix} x_{j_1} \\ x_{j_2} \end{pmatrix} + V'' \begin{pmatrix} x_{j_3} \\ x_{j_4} \end{pmatrix} \right\| \ll \xi^{1/2},$$

hence

$$\begin{pmatrix} x_{j_1} \\ x_{j_2} \end{pmatrix} = \tilde{V} \begin{pmatrix} x_{j_3} \\ x_{j_4} \end{pmatrix} + \mathbf{y}, \quad \tilde{V} = -(V')^{-1} V'', \quad \|\mathbf{y}\| \ll \xi^{1/2} \Delta^{1/2}. \quad (3.11)$$

If we fix x_{j_3} and x_{j_4} , then the previous discussion shows

$$|\{(x_{j_1}, x_{j_2}) \in \mathbb{Z}^2 : |\mathbf{v}_1^T A\mathbf{x}|^2 + |\mathbf{v}_2^T A\mathbf{x}|^2 \leq \xi\}| \ll (1 + \xi^{1/2} \Delta^{1/2})^2. \quad (3.12)$$

On the other hand, substituting (3.11) into the equation $Q(\mathbf{x}) = \ell$, we obtain a *binary* quadratic form \tilde{Q} given by a matrix $\begin{pmatrix} 2\alpha_1 & \alpha_2 \\ \alpha_2 & 2\alpha_3 \end{pmatrix}$, say, such that

$$\tilde{Q}(x_{j_3}, x_{j_4}) = \ell + O(\|\mathbf{y}\|(\|\mathbf{Ax}\| + \|\mathbf{Ay}\|)) = \ell + O(\xi^{1/2} \Delta \ell^{1/2} + \xi \Delta^2). \quad (3.13)$$

Concretely, if P is the permutation matrix sending $(x_{j_1}, x_{j_2}, x_{j_3}, x_{j_4})$ to (x_1, x_2, x_3, x_4) , then the 2-by-2 matrix \tilde{A} of \tilde{Q} is given by

$$\tilde{A} = (I_2 \tilde{V}) P^T A P \begin{pmatrix} I_2 \\ \tilde{V} \end{pmatrix}$$

with $I_2 = \begin{pmatrix} 1 & \\ & 1 \end{pmatrix}$. We proceed to count the number of integer points x_{j_3} and x_{j_4} in the small annulus (3.13). To this end, we approximate \tilde{Q} by a rational quadratic form. By Dirichlet's approximation theorem we find for any $H \geq 1$ three fractions a_1/q , a_2/q , and a_3/q with $q \leq H$ such that $|\alpha_j - a_j/q| \leq (qH^{1/3})^{-1}$. Hence (3.13) implies

$$a_1 x_{j_3}^2 + a_2 x_{j_3} x_{j_4} + a_3 x_{j_4}^2 = \ell q + O\left(\xi^{1/2} \Delta \ell^{1/2} q + \xi \Delta^2 q + \frac{\ell}{H^{1/3}}\right).$$

We choose

$$H := 1 + \left(\frac{\ell}{\Delta^2 \xi}\right)^{3/8}$$

so that the error term in the preceding display becomes

$$E := \xi \Delta^2 + (\xi \Delta^2)^{1/8} \ell^{7/8}.$$

Hence the number of integer pairs (x_{j_3}, x_{j_4}) satisfying (3.13) is at most

$$\ll_{\varepsilon} (\ell H(1+E))^{\varepsilon} (1+E).$$

Combining this with (3.12), we arrive at the first bound of the lemma. The second bound is simple: and in order to bound $r_Q(\ell)$, we fix x_1 and x_2 (for which we have $O(\ell)$ choices); the remaining binary problem has at most ℓ^{ε} solutions. \blacksquare

3.3 Quaternions and quaternary quadratic spaces

The pair $(B(\mathbb{R}), \text{nr})$ is a nondegenerate quadratic space the polarization of which is the inner product

$$(\gamma_1, \gamma_2)_B = \frac{1}{2} \text{tr}(\gamma_1 \overline{\gamma_2}). \quad (3.14)$$

Given an order \mathcal{O} , the choice of a \mathbb{Z} -basis $\{\gamma_1, \dots, \gamma_4\}$ identifies the quadratic lattice (\mathcal{O}, nr) with the integral positive definite quadratic lattice (\mathbb{Z}^4, Q) , the associated matrix of which is

$$A = (\text{tr}(\gamma_i \overline{\gamma_j}))_{1 \leq i, j \leq 4} \in GL_4(\mathbb{Z}),$$

namely

$$\text{nr}(x_1 \gamma_1 + \dots + x_4 \gamma_4) = Q(x_1, \dots, x_4).$$

The choice of a different basis gives a \mathbb{Z} -equivalent quadratic form, and the set of quadratic forms associated with the orders locally conjugate to \mathcal{O} are precisely the forms in the genus of Q (the forms everywhere locally equivalent to Q).

By the definition of the discriminant,

$$\text{disc}(\mathcal{O}) = \det(A).$$

The different and the reduced discriminant of \mathcal{O} admit a similar interpretation in terms of the quadratic form: if $\{\gamma_1, \dots, \gamma_4\}$ forms a \mathbb{Z} -basis of \mathcal{O} , then a basis of the dual lattice \mathcal{O}^* is given by the dual basis $\gamma_1^*, \dots, \gamma_4^*$ with $\text{tr}(\gamma_i^* \overline{\gamma_j^*}) = \delta_{ij}$. It follows that the 4-by-4 matrix $(\text{tr}(\gamma_i^* \overline{\gamma_j^*}))$ is the inverse of A .

If moreover \mathcal{O} is an Eichler order, it follows from the previous remark and (2.2) that the reduced discriminant $\text{disc}^*(\mathcal{O})$ is the smallest number such that multiplication with A^{-1} gives an even integral matrix; in other terms, $\text{disc}^*(\mathcal{O})$ is just the level of the quadratic form Q .

In particular, since $\text{disc}(\mathcal{O}) = \text{disc}^*(\mathcal{O})^2$, Lemma 1 applied to such Q gives

$$\sum_{\ell \in \mathcal{A}} r_{\mathcal{O}}(\ell) \ll_{\varepsilon} \frac{y^2}{\text{disc}(\mathcal{O})^{1/2}} + \frac{y^{3/2}}{\text{disc}(\mathcal{O})^{1/4}} + y^{\varepsilon} |A| \quad (3.15)$$

for $\mathcal{A} \subseteq \mathbb{N} \cap [1, y]$, where

$$r_{\mathcal{O}}(\ell) = |\{\gamma \in \mathcal{O}, \text{nr}(\gamma) = \ell\}|.$$

For any two-dimensional subspace S of $B(\mathbb{R})$ and $\gamma \in (B(\mathbb{R}), \text{nr})$ let γ_S denote the orthogonal projection of γ onto S . We write

$$r_{\mathcal{O}}(\ell, \xi) = \sup_{\dim S=2} |\{\gamma \in \mathcal{O}, \text{nr}(\gamma) = \ell, \text{nr}(\gamma_S) \leq \xi\}|.$$

Then we can rephrase Lemma 2 as

$$r_{\mathcal{O}}(\ell, \xi) \ll_{\varepsilon} (1 + \xi \text{disc}(\mathcal{O})^2)(1 + \text{disc}(\mathcal{O})^2 \xi + (\text{disc}(\mathcal{O})^2 \ell^7 \xi)^{1/8})(\ell \text{disc}(\mathcal{O})(\xi + 1/\xi))^{\varepsilon} \quad (3.16)$$

and

$$r_{\mathcal{O}}(\ell, \xi) \leq r_{\mathcal{O}}(\ell) \ll_{\varepsilon} \ell^{1+\varepsilon}. \quad (3.17)$$

4 Adelic Realization of Spheres

In this section, we realize the space $X(\mathcal{O})$ discussed in the introduction as an adelic quotient associated with B ; first, we recall how to realize the 2-sphere as a homogeneous space for the group $PB^{\times}(\mathbb{R})$.

4.1 Traceless quaternions and ternary forms

Let

$$B^0 = \{\gamma \in B, \text{tr}(\gamma) = 0\}$$

denote the space of trace zero quaternions. Restricting the norm form, (B^0, nr) is a quadratic space (this is the subspace orthogonal to the scalars $Z(B)$). The group of invertible quaternions B^{\times} acts on B^0 by conjugation

$$\gamma \in B^{\times} : x \in B^0 \mapsto \gamma.x := \gamma x \gamma^{-1} \in B^0,$$

and this action is isometric and its kernel is given by the scalar matrices. If we denote by $PB^{\times} := Z(B^{\times}) \backslash B^{\times}$ the projective group of invertible elements of B (considered as an algebraic group acting on B^0 via conjugation), then the above map defines an isomorphism of \mathbb{Q} -algebraic groups [23, Theorem I.3.3]

$$SO(B^0, \text{nr}) \simeq PB^{\times}.$$

In particular, the choice of an orthonormal basis of $B^0(\mathbb{R})$ (for instance $\{i/|a|^{1/2}, j/|b|^{1/2}, ij/|ab|^{1/2}\}$) defines for definite quaternion algebras B an identification

- of (B^0, nr) with three-dimensional Euclidean space,
- of the ellipsoid $S_{\text{nr}}^2 := \{x \in B^0(\mathbb{R}), \text{nr}(x) = 1\}$ with S^2 (with the third basis vector identified with the north pole),
- and of $\text{PB}^\times(\mathbb{R})$ with $\text{SO}_3(\mathbb{R})$.

We fix such an identification and will make no distinction between the real groups or the underlying real vector spaces.

Let $x_\infty \in S_{\text{nr}}^2$ be (say) the third basis vector (north pole) and $K_\infty = \text{PB}^\times(\mathbb{R})_{x_\infty}$ its stabilizer which is generated by $1, x \in B^\times(\mathbb{R})$. Then K_∞ is the rotation group in the plane orthogonal to the line through the north pole and the south pole, hence $K_\infty \simeq S^1$. By Witt's theorem the map

$$g \in \text{PB}^\times(\mathbb{R}) \mapsto g.x_\infty \in S_{\text{nr}}^2 \simeq S^2$$

identifies $\text{PB}^\times(\mathbb{R})/K_\infty$ with the ellipsoid. In particular, we will view functions on S^2 as function on $\text{PB}^\times(\mathbb{R})$ which are left invariant by K_∞ .

4.2 Adelic quaternions

Let \mathbb{A} denote the \mathbb{Q} -algebra of adeles, \mathbb{A}_f the finite adeles, let $B(\mathbb{A}) = B(\mathbb{Q}) \otimes_{\mathbb{Q}} \mathbb{A}$, $B(\mathbb{A}_f) = B(\mathbb{Q}) \otimes_{\mathbb{Q}} \mathbb{A}_f$ be the algebra of adelic points of B , and let $\text{PB}^\times(\mathbb{A})$, $\text{PB}^\times(\mathbb{A}_f)$ be the groups of adelic points of PB^\times . An element x of $B(\mathbb{A})$ will be expressed in terms of its local components as $x = (x_v)_v = x_{\mathbb{R}}(x_p)_p = x_{\mathbb{R}}x_f$. For v a place, x_v will denote either an element of B_v or the element of $B(\mathbb{A})$ the v th component of which is x_v and the other components of which are trivial (the same notation holds for the other adelic spaces). The group of rational points $\text{PB}^\times(\mathbb{Q})$ is embedded diagonally as a discrete cocompact quotient of $\text{PB}^\times(\mathbb{A})$ (since B is a division algebra). We denote this quotient by

$$X_{\text{PB}^\times} := \text{PB}^\times(\mathbb{Q}) \backslash \text{PB}^\times(\mathbb{A}).$$

Given an Eichler order \mathcal{O} , we denote by $\widehat{\mathcal{O}}$ the closure of \mathcal{O} in $B(\mathbb{A}_f)$, and by

$$K_f = \mathbb{A}_f^\times \backslash \mathbb{A}_f^\times \cdot \widehat{\mathcal{O}}^\times \subset \text{PB}^\times(\mathbb{A}_f)$$

the image $\widehat{\mathcal{O}}^\times$ in $\text{PB}^\times(\mathbb{A})$; this is an open compact subgroup of $\text{PB}^\times(\mathbb{A}_f)$. Let

$$X_{\text{PB}^\times}(\mathcal{O}) := \text{PB}^\times(\mathbb{Q}) \backslash \text{PB}^\times(\mathbb{A}) / K_f;$$

this quotient decomposes into a finite union of components indexed by the quotient

$$\mathrm{Cl}(K_f) := \mathrm{PB}^\times(\mathbb{Q}) \backslash \mathrm{PB}^\times(\mathbb{A}_f) / K_f \simeq \mathrm{B}^\times(\mathbb{Q}) \backslash \mathrm{B}^\times(\mathbb{A}_f) / \widehat{\mathcal{O}}^\times$$

(the later identification holds since $\mathbb{A}^\times = \mathbb{Q}^\times \widehat{\mathbb{Z}}^\times$). More precisely, the map

$$\gamma_f = (\gamma_p)_p \in \mathrm{B}^\times(\mathbb{A}_f) \mapsto (\gamma_p \mathcal{O}_p)_p =: (\mathfrak{I}_p) \mapsto \mathfrak{I} := \bigcap_p \mathrm{B}^\times(\mathbb{Q}) \cap \mathfrak{I}_p$$

identifies $\mathrm{Cl}(K_f)$ with $\mathrm{Cl}(\mathcal{O})$, the set of equivalence classes of right- \mathcal{O} ideals (see Section 2.2). Let us choose a set of representatives of such ideal classes (one may choose one to be \mathcal{O}), let us denote by I an ideal in that set (denote by $[I]$ the corresponding ideal class) and let $\gamma_{[I]} \in \mathrm{B}^\times(\mathbb{A}_f)$ be such that $I \cdot \widehat{\mathcal{O}} = \gamma_{[I]} \widehat{\mathcal{O}}$. Since

$$\mathrm{B}^\times(\mathbb{A}) = \bigsqcup_{[I] \in \mathrm{Cl}(\mathcal{O})} \mathrm{B}^\times(\mathbb{Q}) \gamma_{[I]} \widehat{\mathcal{O}}^\times \mathrm{B}^\times(\mathbb{R}),$$

we have a natural identification

$$X_{\mathrm{PB}^\times}(\mathcal{O}) \simeq \bigsqcup_{[I]} \Gamma_{[I]} \backslash \mathrm{PB}^\times(\mathbb{R}), \quad \Gamma_{[I]} = \mathrm{PB}^\times(\mathbb{Q}) \cap \gamma_{[I]} K_f \gamma_{[I]}^{-1} = \mathcal{O}_I(I)^\times / \{\pm 1\}.$$

We obtain the identification

$$X(\mathcal{O}) := X_{\mathrm{PB}^\times}(\mathcal{O}) / K_\infty = \mathrm{PB}^\times(\mathbb{Q}) \backslash \mathrm{PB}^\times(\mathbb{R}) \mathrm{PB}^\times(\mathbb{A}_f) / K_\infty K_f \simeq \bigsqcup_{[I]} \Gamma_{[I]} \backslash S^2. \quad (4.1)$$

In the sequel, we will thus identify the functions on this union of quotients of spheres with the functions on X_{PB^\times} which are invariant under the right multiplication by the elements of $K_\infty K_f$.

We fix the Haar measure on $\mathrm{PB}^\times(\mathbb{A})$ as follows. We fix the Haar measure on $\mathrm{SO}_3(\mathbb{R}) \simeq \mathrm{PB}^\times(\mathbb{R})$, and on K_∞ we take the Haar probability measure; the $\mathrm{SO}_3(\mathbb{R})$ -invariant quotient measure on S^2 is then a probability measure. We fix the Haar measure on $\mathrm{PB}^\times(\mathbb{A}_f)$ which gives K_f measure 1. Since

$$w_{[I]} := |\Gamma_{[I]}| \leq 12$$

for any $[I]$ (see Section 2.2), it follows from (2.4) that

$$\mathrm{vol}(X_{\mathrm{PB}^\times}(\mathcal{O})) = \mathrm{vol}(X(\mathcal{O})) \asymp |\mathrm{Cl}(\mathcal{O})| = \mathrm{disc}(\mathcal{O})^{1/2+o(1)}. \quad (4.2)$$

4.3 The irreducible representations of $\mathrm{SO}_3(\mathbb{R})$

In this section, we recall the relation between Laplace eigenfunctions on the sphere S^2 and the irreducible representations of $\mathrm{SO}_3(\mathbb{R})$ (see [6, Section 7]). As above, we fix an identification of $\mathrm{PB}^\times(\mathbb{R})$ with $\mathrm{SO}_3(\mathbb{R})$ in such a way that x_∞ is identified with the north pole.

An eigenvalue λ of Δ_{S^2} is of the form

$$\lambda = m(m+1)$$

for some integer $m \geq 0$, and the space of λ -eigenforms is precisely the

$$d_{2m} = 2m + 1$$

dimensional space \mathcal{H}_m of *spherical harmonics* of degree m . Recall that spherical harmonics are functions φ which are restrictions to S^2 of a homogeneous harmonic polynomial (a polynomial P in three variables satisfying $\Delta_{\mathbb{R}^3} P = 0$). The group $\mathrm{SO}_3(\mathbb{R})$ acts on \mathcal{H}_m via its action on S^2 , and this defines an irreducible representation π_{2m} , say, of dimension d_{2m} . The representations $(\pi_{2m})_{m \geq 0}$ exhaust the set of isomorphism classes of irreducible representations of $\mathrm{SO}_3(\mathbb{R})$.

Via the identification of S^2 with $\mathrm{SO}_3(\mathbb{R})/K_\infty$ we may (and will) view functions on the sphere with right- K_∞ functions on $\mathrm{SO}_3(\mathbb{R})$. In this identification, the action of the Laplace operator coincides with the restriction of the action of the *Casimir element* $\Omega \in \mathcal{ZU}(\mathfrak{so}_3)$, derived from the right multiplication of $\mathrm{SO}_3(\mathbb{R})$ to the space of right K_∞ -invariant functions (remember that the action of Ω commutes with right K_∞ -translations). In that way, \mathcal{H}_m corresponds to the subspace of K_∞ -invariant Casimir eigenfunctions for the same eigenvalue λ . This identification is just one of the d_{2m} realizations of π_{2m} as an irreducible representation of the (left) regular representation $L^2(\mathrm{SO}_3(\mathbb{R}))$.

Plainly the Casimir element also acts on $L^2(X_{\mathrm{PB}^\times}(\mathcal{O}))$ and we denote by $\mathcal{H}_m(\mathcal{O})$ the Casimir λ eigenspace of K_∞ -invariant functions: this is the subspace from where we choose the function φ of Theorem 1.

4.4 The Hecke operators

The Laplace operator Δ_{S^2} is identified with the restriction of the *Casimir element* to K_∞ -invariant functions and thus is directly related to the action of $\mathrm{PB}^\times(\mathbb{R})$ on itself via

right multiplication. The Hecke operators are defined similarly in terms of the action of $\mathrm{PB}^\times(\mathbb{A}_f)$ on X_{PB^\times} via right multiplication.

Let R denote the right regular representation of $\mathrm{PB}^\times(\mathbb{A})$ on $L^2(\mathrm{PB}^\times(\mathbb{Q})\backslash\mathrm{PB}^\times(\mathbb{A}))$, that is,

$$R(g)(\varphi) = g.\varphi : g' \mapsto \varphi(g'g)$$

for $g \in \mathrm{PB}^\times(\mathbb{A})$, $\varphi \in L^2(\mathrm{PB}^\times(\mathbb{Q})\backslash\mathrm{PB}^\times(\mathbb{A}))$. For a smooth (locally constant) compactly supported function f on $\mathrm{PB}^\times(\mathbb{A}_f)$, one associates the operator

$$R(f) : \varphi \mapsto R(f)(\varphi) = \int_{\mathrm{PB}^\times(\mathbb{A})} f(g)R(g)(\varphi) \, dg.$$

(We follow the standard notation and denote by R both the right regular representation and the following operator; this will not lead to confusion.) The fact that R is unitary implies the following adjointness property:

$$R(f)^* = R(\hat{f}), \quad \hat{f}(g) = \bar{f}(g^{-1}). \quad (4.3)$$

The map $f \mapsto R(f)$ is a representation of the convolution algebra of smooth compactly supported functions on $\mathrm{PB}^\times(\mathbb{A}_f)$:

$$R(f_1 * f_2) = R(f_1) \circ R(f_2) \quad \text{for } f_1 * f_2(h) = \int_{\mathrm{PB}^\times(\mathbb{A}_f)} f_1(g) f_2(g^{-1}h) \, dg.$$

Let $\mathcal{H}(K_f)$ be the subalgebra generated by the characteristic functions of the sets of the form $K_f \gamma_f K_f$ for $\gamma_f = (\gamma_p)_p \in \mathrm{PB}^\times(\mathbb{A}_f)$ such that $\gamma_p = 1$ for p dividing $D_B N$. For such a γ_f we define $R(\mathbf{1}_{K_f \gamma_f K_f})$ to be the Hecke operator associated with γ_f and

$$\mathbf{T} := R(\mathcal{H}(K_f))$$

the algebra of Hecke operators. It is clear that \mathbf{T} maps $L^2(\mathrm{PB}^\times(\mathbb{Q})\backslash\mathrm{PB}^\times(\mathbb{A}))$ onto the space of right K_f -invariant functions $L^2(\mathrm{PB}^\times(\mathbb{Q})\backslash\mathrm{PB}^\times(\mathbb{A}))^{K_f}$. Let us recall how $\mathcal{H}(K_f)$ acts on the latter via correspondences. We have normalized dg so that K_f has a measure 1. For γ_f as above we have a finite disjoint union

$$K_f \gamma_f K_f = \bigsqcup_i \gamma_{f,i} K_f,$$

and for a K_f -invariant function φ we have

$$R(\mathbf{1}_{K_f \gamma_f K_f}) \varphi(x) = \sum_i \varphi(x \gamma_{f,i}).$$

It is easy to see that $K_f \gamma_f^{-1} K_f = K_f \gamma_f K_f$; this implies that $R(\mathbf{1}_{K_f \gamma_f K_f}) = R(\mathbf{1}_{K_f \gamma_f^{-1} K_f})$ is self-adjoint. The Hecke algebra $\mathcal{H}(K_f)$ is commutative and is generated by the functions

$$f_p := \frac{1}{p^{1/2}} \mathbf{1}_{K_f \gamma_p K_f}$$

for p varying over the primes coprime with $\text{disc}(\mathcal{O})$ and $\gamma_p \in B_p$ any element of norm $p^{\pm 1}$ (the function f_p does not depend on the choice of the element γ_p , for instance, one could take $\gamma_p = \begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix}_p$). More generally, we define

$$f_{p^\alpha} := \frac{1}{p^{\alpha/2}} \mathbf{1}_{K_f \begin{pmatrix} p^\alpha & 0 \\ 0 & 1 \end{pmatrix}_p K_f}, \quad T_{p^\alpha} = R(f_{p^\alpha}), \quad (4.4)$$

and for $n = p_1^{\alpha_1} \dots p_r^{\alpha_r}$ coprime with $\text{disc}(\mathcal{O})$ we write

$$f_n = f_{p_1^{\alpha_1}} * \dots * f_{p_r^{\alpha_r}}, \quad T_n := R(f_n) = T_{p_1^{\alpha_1}} \dots T_{p_r^{\alpha_r}}.$$

The f_n and T_n satisfy the usual multiplicativity relations of Hecke operators:

$$f_m * f_n = \sum_{d|(m,n)} f_{mn/d^2}, \quad T_m T_n = \sum_{d|(m,n)} T_{mn/d^2}. \quad (4.5)$$

As is well known, the Casimir operator may be diagonalized in an orthonormal basis of K_∞ -finite eigenvectors. Since the actions of $\text{PB}^\times(\mathbb{A}_f)$ and $\text{PB}^\times(\mathbb{R})$ by right translation on $\text{PB}^\times(\mathbb{Q}) \backslash \text{PB}^\times(\mathbb{A})$ commute, the Hecke operators also commute with the Casimir operator, and $L^2(X_{\text{PB}^\times}(\mathcal{O}))$ admits an orthonormal basis of (\mathbf{T}, Ω) -eigenfunctions containing an orthonormal basis of $\mathcal{H}_m(\mathcal{O})$. For φ a Hecke eigenfunction, we denote by λ_φ the corresponding character of the Hecke algebra: for $f \in \mathcal{H}(K_f)$ one has

$$R(f)\varphi = \lambda_\varphi(f)\varphi.$$

In particular, by the multiplicativity relation (4.5), one has

$$\lambda_\varphi(f_p)^2 - \lambda_\varphi(f_{p^2}) = 1 \quad \text{which implies that } |\lambda_\varphi(f_p)| + |\lambda_\varphi(f_{p^2})| \gg 1. \quad (4.6)$$

By Jacquet–Langlands correspondence [13], there exists a holomorphic Hecke-eigenform $f(z)$ on a quotient of the upper half plane such that for any prime p coprime

with $\text{disc}(\mathcal{O})$ one has $\lambda_\varphi(f_{\mathfrak{p}^\alpha}) = \lambda_f(\mathfrak{p}^\alpha)$. Although we will not need it, we remark that from Deligne's bound on the Hecke eigenvalues of such holomorphic modular forms one has sharp control on the size of this character: for instance, for φ a nonconstant Hecke-eigenfunction one has $|\lambda_\varphi(f_{\mathfrak{p}^\alpha})| \leq (\alpha + 1)$.

5 The Pre-trace Formula

5.1 Automorphic kernel

The starting point of the proof follows closely [11]: given an orthonormal basis $\{\varphi_{-m}, \dots, \varphi_m\}$ of $\mathcal{H}_m \subset L^2(\text{PB}^\times(\mathbb{R}))$, consider the “reproducing” kernel

$$K_m(g, h) = \sum_i \varphi_i(g) \overline{\varphi_i(h)}, \quad g, h \in \text{PB}^\times(\mathbb{R}).$$

The associated integral operator

$$\varphi \mapsto P_m(g) = \int_{\text{PB}^\times(\mathbb{R})} K_m(g, h) \varphi(h) \, dh$$

is the orthogonal projection on \mathcal{H}_m . In particular, this kernel is independent of the choice of the orthonormal basis, and one has

$$K_m(\gamma g, \gamma h) = K_m(g, h) \tag{5.1}$$

for any $\gamma \in \text{PB}^\times(\mathbb{R})$. One has the following identity [6, Proposition 9.4.2]

$$K_m(g, h) = d_m p_m((g.x_\infty, h.x_\infty)), \tag{5.2}$$

where $(\cdot, \cdot) = (\cdot, \cdot)_{\text{B}^0(\mathbb{R})} = (\cdot, \cdot)_{\text{B}(\mathbb{R})}$ is the inner product (3.14) on $\text{B}^0(\mathbb{R}) \subset \text{B}(\mathbb{R})$ and p_m is the spherical polynomial of degree m (that is, the m th Legendre polynomial).

Given a smooth compactly supported function f on $\text{PB}^\times(\mathbb{A}_f)$, we consider the automorphic kernel

$$K_{m,f}(x, y) := \sum_{\gamma \in \text{PB}^\times(\mathbb{Q})} f(x_f^{-1} \gamma y_f) K_m(x_{\mathbb{R}}, \gamma y_{\mathbb{R}}), \quad x, y \in \text{PB}^\times(\mathbb{A}).$$

Note that for any given $x, y \in \text{PB}^\times(\mathbb{A})$ the above sum is finite, and moreover $K_{m,f}$ is left $\text{PB}^\times(\mathbb{Q}) \times \text{PB}^\times(\mathbb{Q})$ -invariant: for $\gamma_1 \in \text{PB}^\times(\mathbb{Q})$ one has

$$\begin{aligned} K_{m,f}(\gamma_1 x, y) &= \sum_{\gamma \in \text{PB}^\times(\mathbb{Q})} f(x_f^{-1} \gamma_1^{-1} \gamma y_f) K_m(\gamma_1 x_{\mathbb{R}}, \gamma y_{\mathbb{R}}) \\ &= \sum_{\gamma \in \text{PB}^\times(\mathbb{Q})} f(x_f^{-1} \gamma y_f) K_m(\gamma_1 x_{\mathbb{R}}, \gamma_1 \gamma y_{\mathbb{R}}) = \sum_{\gamma \in \text{PB}^\times(\mathbb{Q})} f(x_f^{-1} \gamma y_f) K_m(x_{\mathbb{R}}, \gamma y_{\mathbb{R}}) \end{aligned}$$

by (5.1). Now let $P_{m,f}$ be the associated integral operator on $L^2(X_{\text{PB}^\times})$,

$$P_{m,f}(\varphi)(x) = \int_{X_{\text{PB}^\times}} K_{m,f}(x, y) \varphi(y) \, dy.$$

By the previous remark this is well defined. Moreover, if $f \in \mathcal{H}(K_f)$ then

$$P_{m,f}(\varphi) \in \mathcal{H}_m(\mathcal{O})$$

as defined at the end of Section 4.3. To see this, we observe that on the one hand, $P_{m,f}(\varphi)$ is right K_f -invariant, since f is left K_f -invariant, hence $P_{m,f}(\varphi) \in L^2(X_{\text{PB}^\times}(\mathcal{O}))$; on the other hand, if φ is orthogonal to $\mathcal{H}_m(\mathcal{O})$, then for any $y_f \in \text{PB}^\times(\mathbb{A}_f)$ the function $x_{\mathbb{R}} \mapsto \varphi(x_{\mathbb{R}} y_f)$ is orthogonal to \mathcal{H}_m , so it is annihilated by P_m ; therefore

$$\begin{aligned} P_{m,f}(\varphi)(x_f x_{\mathbb{R}}) &= \int_{\text{PB}^\times(\mathbb{A})} f(x_f^{-1} y_f) K_m(x_{\mathbb{R}}, y_{\mathbb{R}}) \varphi(y_{\mathbb{R}} y_f) \, dy_f \, dy_{\mathbb{R}} \\ &= \int_{\text{PB}^\times(\mathbb{A}_f)} f(x_f^{-1} y_f) P_m(y_f \cdot \varphi)(x_{\mathbb{R}}) \, dy_f = 0 \end{aligned}$$

as desired.

Let now $\{\varphi_j\}_j$ be an orthonormal Hecke eigenbasis of $\mathcal{H}_m(\mathcal{O})$ which contains our preferred form φ (which we assume to be L^2 -normalized); it follows from the spectral theorem and the previous computation that

$$K_{m,f}(x, y) = \sum_j \lambda_{\varphi_j}(f) \varphi_j(x) \overline{\varphi_j(y)}.$$

We will choose f of the form $\tilde{f} * \hat{\tilde{f}}$ for $\tilde{f} \in \mathcal{H}(K_f)$, cf. (4.3). It then follows that

$$\lambda_{\varphi_j}(f) = |\lambda_{\varphi_j}(\tilde{f})|^2,$$

and hence by positivity

$$|\lambda_\varphi(\tilde{f})|^2 |\varphi(x)|^2 \leq \sum_j \lambda_{\varphi_j}(f) |\varphi_j(x)|^2 = \sum_{\gamma \in \text{PB}^\times(\mathbb{Q})} f(x_f^{-1} \gamma x_f) K_m(x_{\mathbb{R}}, \gamma x_{\mathbb{R}}) \quad (5.3)$$

for any $x \in X(\mathcal{O})$.

5.2 Amplification

The right-hand side of (5.3) is a hermitian form in \tilde{f} and the objective will be to choose \tilde{f} appropriately (in terms of the preferred form φ) so as to minimize the right-hand side while maximizing the diagonal hermitian form $|\lambda_\varphi(\tilde{f})|^2$: this is the principle of the *amplification method* invented by Iwaniec and used in [11] and in many other places. Our choice of \tilde{f} will be a slight variant of the usual choice of an amplifier following [22, Section 4.1]. Specifically, for $L \geq 2$ a parameter to be chosen later and f_n as in (4.4), define

$$a_\ell = \begin{cases} \text{sign}(\lambda_\varphi(f_\ell)) & \text{if } \ell \text{ is a prime in } [L, 2L], \text{ coprime with } \text{disc}(\mathcal{O}), \\ \text{sign}(\lambda_\varphi(f_\ell)) & \text{if } \ell = p^2, \text{ } p \text{ a prime in } [L, 2L], \text{ coprime with } \text{disc}(\mathcal{O}), \\ 0 & \text{otherwise,} \end{cases}$$

and let

$$\tilde{f} = \sum_{\ell \leq 4L^2} a_\ell f_\ell$$

so that

$$R(\tilde{f}) = \sum_{\ell \leq 4L^2} a_\ell T_\ell.$$

By construction and (4.6),

$$\lambda_\varphi(\tilde{f}) \gg L / \log L$$

as soon as L is greater than a fixed positive power of $\text{disc}(\mathcal{O})$. Hence (5.3) implies

$$|\varphi(x)|^2 \ll \frac{\log^2 L}{L^2} \sum_{\gamma \in \text{PB}^\times(\mathbb{Q})} f(x_f^{-1} \gamma x_f) K_m(x_{\mathbb{R}}, \gamma x_{\mathbb{R}}). \quad (5.4)$$

We now proceed to bound the γ -sum. From the Hecke multiplicativity relation (4.5), one has

$$f = \tilde{f} * \hat{\tilde{f}} = \sum_d \sum_{\ell_1, \ell_2} a_{d\ell_1} a_{d\ell_2} f_{\ell_1 \ell_2},$$

and we need then to evaluate

$$L^{\varepsilon-2} \left| \sum_d \sum_{\ell_1, \ell_2} a_{d\ell_1} a_{d\ell_2} \sum_{\gamma \in \text{PB}^\times(\mathbb{Q})} f_{\ell_1 \ell_2}(x_f^{-1} \gamma x_f) K_m(x_{\mathbb{R}}, \gamma x_{\mathbb{R}}) \right|.$$

Let $\mathcal{O}' \subset \text{B}(\mathbb{Q})$ be the unique order the closure in $\text{B}(\mathbb{A}_f)$ of which is $x_f \widehat{\mathcal{O}} x_f^{-1}$:

$$\mathcal{O}' = \bigcap_p \text{B}(\mathbb{Q}) \cap x_{f,p} \mathcal{O}_p x_{f,p}^{-1}.$$

This is an order everywhere locally conjugate to \mathcal{O} , thus is an Eichler order associated with the integer N . Consider a $\gamma \in \text{PB}^\times(\mathbb{Q})$ that contributes non trivially to the above sum. This implies that there is $\gamma \in \text{B}^\times(\mathbb{Q})$ such that

$$\gamma \in \mathcal{O}'_p \times x_{f,p} \begin{pmatrix} p^{v_p(\ell_1 \ell_2)} & 0 \\ 0 & 1 \end{pmatrix} x_{f,p}^{-1} \mathcal{O}'_p \times \subset \mathcal{O}'_p, \quad (p, \text{disc}(\mathcal{O})) = 1,$$

and $\gamma \in \mathcal{O}'_p \times$ if $p \mid \text{disc}(\mathcal{O})$. Hence γ is contained in \mathcal{O}' and $\text{nr}(\gamma) = \ell_1 \ell_2$. From this, (5.2) and (5.4), we obtain

$$|\varphi(x)|^2 \ll \frac{d_m}{L^{2-\varepsilon}} \sum_{d, \ell_1, \ell_2} \frac{|a_{d\ell_1}| |a_{d\ell_2}|}{\sqrt{\ell_1 \ell_2}} \left| \sum_{\substack{\gamma \in \mathcal{O}' \\ \text{nr}(\gamma) = \ell_1 \ell_2}} p_m((x', \gamma \cdot x')) \right| \quad (5.5)$$

where $x' := x_{\mathbb{R}} x_{\infty} x_{\mathbb{R}}^{(-1)} \in \text{B}^\times(\mathbb{R})$.

Remark. The sum

$$\sum_{\substack{\gamma \in \mathcal{O}' \\ \text{nr}(\gamma) = \ell_1 \ell_2}} p_m((x', \gamma \cdot x')) \quad (5.6)$$

is the $\ell_1 \ell_2$ th Fourier coefficient of a theta series associated with the quaternary quadratic lattice $(\mathcal{O}', \text{nr})$ (which by definition is contained in the genus of (\mathcal{O}, nr)). The occurrence of this theta series should not be a surprise: more precise computations show that the kernel $K_{m,f}(x, y)$ is indeed a linear combination of Fourier coefficients of an (adelic) quaternary theta series. Such computations go back at least to Eichler and yield to the so-called *Shimizu*-theta correspondence between (some) automorphic forms and some automorphic forms associated to the orthogonal groups $\text{SO}(\text{B})$.

This correspondence provides an alternative realization of the Jacquet–Langlands correspondence discussed at the end of the previous section [20]. \square

5.3 Average bounds for Fourier coefficients of theta series

Our main result will follow from bounds for the Fourier coefficients (5.6) on average; for this we use the results of Section 3. The arguments given there are soft and geometric in essence and in particular make no explicit use of the fact that the underlying theta series is modular.

We start by noting the following estimate for the Legendre polynomial [7, 8.918.4]:

$$|p_m(t)| \leq \min\left(1, \frac{1}{m^{1/2}(1-t^2)^{1/4}}\right), \quad t \in [-1, 1], \quad (5.7)$$

with the usual convention $\min(A, \infty) = A$. By definition of the inner product on $(B(\mathbb{R}), \text{nr})$, one has

$$(g \cdot x', h \cdot x') = (gx'g^{-1}, hx'h^{-1})_{B(\mathbb{R})} = \frac{1}{2} \text{tr}(gx'g^{-1} \overline{hx'h^{-1}})$$

for $g, h, x' \in B^\times(\mathbb{R})$. Using this, a straightforward matrix calculation in $B^\times(\mathbb{R})$ gives for x' of norm 1 and $\gamma \in B^\times(\mathbb{R})$

$$(x', \gamma \cdot x') = (x', \gamma x' \gamma^{-1}) = -1 + 2 \frac{(\gamma, 1)^2 + (\gamma, x')^2}{\text{nr}(\gamma)} = -1 + 2 \frac{\text{nr}(\gamma_S)}{\text{nr}(\gamma)} = 1 - 2 \frac{\text{nr}(\gamma_{S^\perp})}{\text{nr}(\gamma)}, \quad (5.8)$$

where γ_S (resp. γ_{S^\perp}) denote the orthogonal projection of γ to the subspace $S := \mathbb{R} + \mathbb{R}x'$ (resp. to the orthogonal subspace S^\perp).

Plugging (5.7) and (5.8) into (5.5) and using that $d_m \asymp m+1$, we obtain

$$\begin{aligned} |\varphi(x)|^2 &\ll_\varepsilon \frac{(m+1)^{1/2}}{L^{2-\varepsilon}} \sum_{d, \ell_1, \ell_2} \frac{|a_{d\ell_1}| |a_{d\ell_2}|}{\sqrt{\ell_1 \ell_2}} \sum_{\substack{\gamma \in \mathcal{O}' \\ \text{nr}(\gamma) = \ell_1 \ell_2}} \min\left((m+1)^{1/2}, \left(\frac{\text{nr}(\gamma)}{\text{nr}(\gamma_S)} + \frac{\text{nr}(\gamma)}{\text{nr}(\gamma_{S^\perp})}\right)^{1/4}\right) \\ &\ll_\varepsilon \sup_{\dim S=2} \frac{(m+1)^{1/2}}{L^{2-\varepsilon}} \sum_{d, \ell_1, \ell_2} \frac{|a_{d\ell_1}| |a_{d\ell_2}|}{\sqrt{\ell_1 \ell_2}} \sum_{\substack{\gamma \in \mathcal{O}' \\ \text{nr}(\gamma) = \ell_1 \ell_2}} \min\left((m+1)^{1/2}, \frac{\text{nr}(\gamma)^{1/4}}{\text{nr}(\gamma_S)^{1/4}}\right) =: S, \end{aligned}$$

say. Since a_ℓ is supported on primes or squares of primes of size $[L, 2L]$, the sum S splits into five parts, say

$$S = \sum_{\alpha=0}^4 \mathcal{S}_\alpha,$$

corresponding to the contribution of the ℓ_1 and ℓ_2 with $\ell_1 \ell_2 \asymp L^\alpha$. We have trivially

$$\mathcal{S}_0 \ll \frac{m+1}{L^{1-\varepsilon}}.$$

Let us now assume $\alpha \geq 1$. For fixed ℓ , there are at most four triplets (ℓ_1, ℓ_2, d) with $a_{d\ell_1} a_{d\ell_2} \neq 0$ and $\ell_1 \ell_2 = \ell$. Let $\mathcal{L}_\alpha := \{\ell_1 \ell_2 \mid \ell_1 \ell_2 \asymp L^\alpha, a_{\ell_1} a_{\ell_2} \neq 0\}$, so that

$$\mathcal{L}_\alpha \subseteq [L^\alpha, (2L)^\alpha] \quad \text{and} \quad |\mathcal{L}_\alpha| \ll L^{\min(\alpha, 2)}. \quad (5.9)$$

With this notation, we have

$$\mathcal{S}_\alpha \ll_\varepsilon \sup_{\dim S=2} \frac{(m+1)^{1/2}}{L^{2+\frac{\alpha}{2}-\varepsilon}} \sum_{\ell \in \mathcal{L}_\alpha} \sum_{\substack{\gamma \in \mathcal{O}' \\ \text{nr}(\gamma)=\ell}} \min \left((m+1)^{1/2}, \frac{\text{nr}(\gamma)^{1/4}}{\text{nr}(\gamma_S)^{1/4}} \right). \quad (5.10)$$

We are now in a position to apply the results of Section 3.3. We write

$$V := \text{disc}(\mathcal{O}')^{1/2} = \text{disc}(\mathcal{O})^{1/2} = \text{vol}(X(\mathcal{O}))^{1+o(1)}, \quad (5.11)$$

(cf. (4.2)), and

$$t := (1 + \lambda)^{1/2} \asymp m + 1.$$

5.4 Bound in the volume aspect

In this section, we use only the trivial bound for the Legendre polynomial and estimate the minimum in (5.10) by its first term. We have by (3.15), (5.9), and (5.11)

$$\begin{aligned} \mathcal{S}_1 &\ll_\varepsilon \frac{t}{L^{5/2}} \left(\frac{L^2}{V} + \frac{L^{3/2}}{V^{1/2}} + L \right) (LV)^\varepsilon, \\ \mathcal{S}_\alpha &\ll_\varepsilon \frac{t}{L^{2+\frac{\alpha}{2}}} \left(\frac{L^{2\alpha}}{V} + \frac{L^{\frac{3}{2}\alpha}}{V^{1/2}} + L^2 \right) (LV)^\varepsilon \quad (2 \leq \alpha \leq 4), \end{aligned}$$

and hence

$$S \ll_{\varepsilon} t \left(\frac{1}{L} + \frac{L^4}{V} + \frac{L^2}{V^{1/2}} \right) (LV)^{\varepsilon}.$$

Choosing $L = V^{-1/6}$ we find

$$\varphi(x) \ll_{\varepsilon} t^{1/2} V^{-1/12+\varepsilon}. \quad (5.12)$$

5.5 Bound in the λ -aspect

Using (3.16), we first estimate the contribution of those γ in the innermost sum of (5.10) where $\text{nr}(\gamma_S) \leq (LVt)^{-100}$ by

$$\ll \frac{t^{1/2}}{L^{2+\frac{\alpha}{2}}} L^{\min(\alpha, 2)} t^{1/2} (LVt)^{\varepsilon} \ll \frac{t}{L} (LVt)^{\varepsilon}.$$

We cut the remaining sum over γ into $O((LVt)^{\varepsilon})$ dyadic pieces $\text{nr}(\gamma_S) \asymp \xi$ where $(LVt)^{-100} \leq \xi \leq L^4$. Combining (3.16), (3.17), (5.9), and (5.10), we find

$$\begin{aligned} S_{\alpha} \ll (LVt)^{\varepsilon} & \left(\frac{t}{L} + \max_{\xi} \frac{t^{1/2}}{L^{2+\frac{\alpha}{2}}} L^{\min(\alpha, 2)} \min \left(t^{1/2}, \frac{L^{\alpha/4}}{\xi^{1/4}} \right) \right. \\ & \left. \times \min(L^{\alpha}, (1 + \xi V^2)(1 + \xi V^4 + \xi^{1/8} L^{7\alpha/8} V^{1/2})) \right). \end{aligned}$$

There are only two possible choices for ξ where the preceding expression can become maximal, namely

$$\xi = \frac{L^{\alpha}}{t^2}$$

or the solution to $(1 + \xi V^2)(1 + \xi V^4 + \xi^{1/8} L^{7\alpha/8} V^{1/2}) = L^{\alpha}$ which is of size

$$\xi \asymp \frac{1}{V^2} \min \left(\frac{L^{\alpha}}{V^2}, \left(\frac{L^{\alpha}}{V^2} \right)^{1/9} \right).$$

Substituting these two values, we obtain

$$S \ll (LVt)^{\varepsilon} \left(\frac{t}{L} + t^{1/2} L^2 V + t^{1/2} L^{26/9} V^{5/9} + \frac{L^2 V^4}{t} + L^2 V^{1/2} t^{3/4} + \frac{L^6 V^6}{t^3} + \frac{L^6 V^{5/2}}{t^{5/4}} \right).$$

Let us now assume

$$t \geq V^4$$

and choose $L = t^{1/12} V^{-1/6} \geq 1$. Then the above expression simplifies to

$$S \ll t^{11/12} V^{1/6} (tV)^\varepsilon$$

and hence

$$\varphi(x) \ll t^{11/24} V^{1/12} (tV)^\varepsilon. \quad (5.13)$$

If $t \leq V^4$, this remains true by (5.12).

Theorem 1 follows now easily from (5.12) and (5.13).

6 Connections with values of L -functions

In this section, we discuss the implications of Theorem 1 to the *subconvexity problem* for automorphic L -functions. This is merely intended as an interesting remark, and in particular we do not provide all details for the facts described here.

Let us recall that the subconvexity problem consists in providing nontrivial bounds for L -values $L(\pi, s)$, where π is an automorphic representation and s is on the critical line ($s = \frac{1}{2} + it$, $t \in \mathbb{R}$). The trivial bound (or convexity bound) is

$$L(\pi, s) \ll C(\pi, s)^{1/4+o(1)}$$

for $C(\pi, s)$ the *analytic conductor* of $L(\pi, s)$; we refer to [10] for its definition. The subconvexity problem consists in lowering the exponent $\frac{1}{4}$. The possibility of replacing $\frac{1}{4}$ by 0 is a form of the Generalized Lindelöf hypothesis.

We consider the following situation. Let K be an imaginary quadratic field which admits a (\mathbb{Q} -algebra) embedding $\iota: K \hookrightarrow \mathbb{B}$. To such a pair (K, ι) corresponds a finite set of CM -points

$$CM(K, \iota) \subset X(\mathcal{O}),$$

and we may evaluate our preferred eigenfunction φ along this distinguished set of points. To simplify things slightly, we will suppose that \mathcal{O} is a maximal order and that $\iota(K) \cap \mathcal{O} = \iota(\mathcal{O}_K)$ where \mathcal{O}_K is the ring of integers of K (in other terms, ι defines an optimal embedding of \mathcal{O}_K into \mathcal{O}). In that case the $CM(K, \iota)$ admits an action by the ideal class group $\text{Cl}(\mathcal{O}_K)$ and this action is almost transitive (the number of orbits is 2^r where r is the number of (finite) primes at which \mathbb{B} is ramified) [1]. One has then a beautiful formula due to Waldspurger [24] (see also [8, 10]) which computes the (twisted by χ) average of φ

along $CM(K, \iota)$: given $z \in CM(K, \iota)$, χ a character of $\text{Cl}(\mathcal{O}_K)$ and φ a Laplace–Hecke eigenfunction (more precisely, φ should correspond to a factorizable vector in the associated automorphic representation of $\text{PB}^\times(\mathbb{Q}) \backslash \text{PB}^\times(\mathbb{A})$) with eigenvalue λ , one has

$$\left| \frac{1}{|\text{Cl}(\mathcal{O}_K)|} \sum_{\sigma \in \text{Cl}(\mathcal{O}_K)} \chi(\sigma) \varphi(\sigma.z) \right|^2 = c(\varphi, K) \frac{L(\pi \otimes \pi_\chi, \frac{1}{2})}{\text{disc}(\mathcal{O}_K)^{1/2} \text{disc}(\mathcal{O})^{1/2}} \|\varphi\|_2^2. \quad (6.1)$$

Here $c(\varphi, \chi)$ is positive and satisfies

$$c(\varphi, \chi) = ((1 + \lambda) \text{disc}(\mathcal{O}_K) \text{disc}(\mathcal{O}))^{o(1)};$$

π is a GL_2 -automorphic cuspidal representation of conductor $q_\pi = D_B = \text{disc}(\mathcal{O})^{1/2}$ which is a holomorphic discrete series at ∞ of weight $m + 2 \asymp (1 + \lambda)^{1/2}$ and which has the same Hecke eigenvalues as φ for the Hecke operators T_p at the primes $p \nmid \text{disc}(\mathcal{O})$; π_χ is the GL_2 -automorphic representation attached to χ by quadratic base change (or via the theta correspondence) with conductor $q_\chi = \text{disc}(\mathcal{O}_K)$; finally, $L(\pi \otimes \pi_\chi, s)$ is the Rankin–Selberg L -function attached to the pair (π, π_χ) . The analytic conductor of this L -function satisfies

$$C(\pi \otimes \pi_\chi, \frac{1}{2}) \asymp q_{\pi_\chi}^2 q_\pi^2 ((1 + \lambda)^{1/2})^4 \asymp \text{disc}(\mathcal{O}_K)^2 \text{disc}(\mathcal{O}) (1 + \lambda)^2$$

so the *convexity* bound for this L -values is given by

$$L(\pi \otimes \pi_\chi, \frac{1}{2}) \ll \text{disc}(\mathcal{O}_K)^{1/2+o(1)} \text{disc}(\mathcal{O})^{1/4+o(1)} (1 + \lambda)^{1/2+o(1)}.$$

Now (6.1) and Theorem 1 give

$$L(\pi \otimes \pi_\chi, \frac{1}{2}) \ll \text{disc}(\mathcal{O}_K)^{1/2+o(1)} \text{disc}(\mathcal{O})^{1/2-\delta} (1 + \lambda)^{1/2-\delta}.$$

Thus, Theorem 1 yields a subconvex bound in the λ -aspect whatever the value of $\delta > 0$ is; on the other hand, in order to have a subconvex bound in the q_π -aspect, one would need $\delta > \frac{1}{4}$. Moreover, the possibility of taking δ arbitrary close to $\frac{1}{2}$ in Theorem 1 would imply the Lindelöf hypothesis.

A subconvex bound for Rankin–Selberg L -functions $L(\pi_1 \otimes \pi_2, \frac{1}{2})$ in terms of the Archimedean parameter of π was first obtained by Sarnak [19] in connection with quantum unique ergodicity. The numerical value has been improved subsequently by many authors, and the best result to date is due to Lau–Liu–Ye. It is interesting that a purely

diophantine method as in the present paper can re-produce a subconvex estimate in this situation. Finally, we note that, by [15, Theorem 1], a uniform subconvex bound holds:

$$L(\pi \otimes \pi_{\chi}, \tfrac{1}{2}) \ll (\text{disc}(\mathcal{O}_K)^{1/2} \text{disc}(\mathcal{O})^{1/4} (1 + \lambda)^{1/2})^{1-\eta},$$

for some absolute constant $\eta > 0$.

Acknowledgements

The first author would like to acknowledge the hospitality and the excellent working conditions at the Ecole Polytechnique Fédérale de Lausanne (EPFL). The authors would also like to thank their referee for a very careful reading of the manuscript.

Funding

This work is supported by a Volkswagen Lichtenberg grant and in part by a European Research Council (ERC) starting grant 258713 (to V.B.) and is partially supported by the ERC advanced research grant number 228304 and the SNF grant number 200021-12529 (to P.M.).

References

- [1] Bertolini, M. and H. Darmon. "Heegner points on mumford-tate curves." *Inventiones Mathematicae* 126, no. 3 (1996): 413–56.
- [2] Blomer, V. and R. Holowinsky. "Bounding sup-norms of cusp forms of large level." *Inventiones Mathematicae* 179, no. 3 (2010): 645–81.
- [3] Cassels, J. W. S. *Rational Quadratic Forms*. London Mathematical Society Monographs 13. London: Academic Press Inc. [Harcourt Brace Jovanovich Publishers], 1978.
- [4] Donnelly, H. "Bounds for eigenfunctions of the Laplacian on compact Riemannian manifolds." *Journal of Functional Analysis* 187, no. 1 (2001): 247–61.
- [5] Eichler, M. "Zur Zahlentheorie der Quaternionen-Algebren." *Journal für die Reine und Angewandte Mathematik* 195 (1955): 127–51 (1956).
- [6] Faraut, J. *Analysis on Lie Groups*. An Introduction. Cambridge Studies in Advanced Mathematics 110. Cambridge: Cambridge University Press, 2008.
- [7] Gradshteyn, I. S. and I. M. Ryzhik. *Table of Integrals, Series, and Products*. Boston, MA: Academic Press Inc., translated from the fourth Russian edition, 1994 (Translation edited and with a preface by Alan Jeffrey).
- [8] Gross, B.H. "Heights and the special values of L -series." *Number theory*, CMS Conference Proceedings 7. Providence, RI: American Mathematical Society, 1987, pp. 115–87.
- [9] Harcos, G. and N. Templier. "On the sup-norm of Maass cusp forms of large level. ii." preprint.

- [10] Hatcher, R. L. "Heights, and L -series." *Canadian Journal of Mathematics* 42, no. 3 (1990): 533–60.
- [11] Iwaniec, H. and P. Sarnak. " L^∞ norms of eigenfunctions of arithmetic surfaces." *Annals of Mathematics* (2) 141, no. 2 (1995): 301–20.
- [12] Iwaniec, H. and P. Sarnak. "Perspectives on the analytic theory of L -functions." *Geometric and Functional Analysis*. Basel: Birkhauser, 2000.
- [13] Jacquet, H. and R. P. Langlands. *Automorphic Forms on $GL(2)$* . Lecture Notes in Mathematics 114. Berlin: Springer, 1970.
- [14] Jorgenson, J. and J. Kramer. "Bounding the sup-norm of automorphic forms." *Geometric and Functional Analysis* 14, no. 6 (2004): 1267–77.
- [15] Michel, P. and A. Venkatesh. "The subconvexity problem for GL_2 ." *Publications Mathématiques. Institut de Hautes Études Scientifiques* 111, no. 1 (2010): 171–271.
- [16] Milićević, D. "Large values of eigenfunctions on arithmetic hyperbolic surfaces." *Duke Mathematical Journal* 155, no. 2 (2010): 365–401.
- [17] Rhodes, J. A. "Sequences of metrics on compact Riemann surfaces." *Duke Mathematical Journal* 72, no. 3 (1993): 725–38.
- [18] Sarnak, P. "Letter to Morawetz." <http://www.math.princeton.edu/sarnak>.
- [19] Sarnak, P. "Estimates for Rankin–Selberg L -functions and quantum unique ergodicity." *Journal of Functional Analysis* 184, no. 2 (2001): 419–53.
- [20] Shimizu, H. "Theta series and automorphic forms on GL_2 ." *Journal of Mathematical Society of Japan* 24, no. 4 (1972): 638–83.
- [21] Templier, N. "On the sup-norm of Maass cusp forms of large level." *Selecta Mathematica* 16, no. 3 (2010): 501–31.
- [22] Venkatesh, A. "Sparse equidistribution problems, period bounds, and subconvexity." *Annals of Mathematics* 172, no. 2 (2010): 989–1094.
- [23] Vignéras, M.-F. *Arithmétique des Algèbres de Quaternions*. Lecture Notes in Mathematics 800. Berlin: Springer, 1980.
- [24] Waldspurger, J.-L. "Sur les valeurs de certaines fonctions L automorphes en leur centre de symétrie." *Compositio Math.* 54, no. 2 (1985): 173–242.